(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 February 2007 (22.02.2007)

PCT

(10) International Publication Number
WO 2007/022107 A2

(51) International Patent Classification:
G06F 7/00 (2006.01)

(21) International Application Number:
PCT/US2006/031691

(22) International Filing Date: 14 August 2006 (14.08.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/707,992 12 August 2005 (12.08.2005) US

(71) Applicant *(for all designated States except US)*: **CORPORATION FOR NATIONAL RESEARCH INITIATIVES** [US/US]; 1895 Preston White Drive, Reston, Virginia 20191 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **BLANCHI, Christophe G.** [US/US]; 49 Valerian Court, Rockville, Maryland 20852 (US). **MORALES, Henry N. Jerez** [BO/US]; 12713 Colony Place Ne, Albuquerque, New Mexico 87122 (US). **LANNOM, Laurence** [US/US]; 4127 Leland Street, Chevy Chase, Maryland 20815 (US). **MANEPALLI, Girdhar** [IN/US]; 12043 Greywing Square, Apt. 3, Reston, Virginia 20191 (US).

(74) Agents: **FEIGENBAUM, David L.** et al.; Fish & Richardson P.C., P.o. Box 1022, Minneapolis, MN 55440 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: MANAGING AND USING SHARED DIGITAL INFORMATION ON A NETWORK

(57) Abstract: Among other things, use of and access to shared information on a network is managed by validating that a putative persistent identifier asserted to be associated with a user who is requesting to use information on a network is actually associated with the user, resolving the persistent identifier to state information about the user and a group to which the user belongs and that is associated with the identifier, determining whether the state information for the group contains the identifier associated with the user, and allowing access and/or use of the shared information based on the permissions allowed for the group.

WO 2007/022107 A2

Managing and Using Shared Digital Information on a Network

This application is entitled to the benefit of the filing date of United States provisional application 60/707,992, filed August 12, 2005, the entire contents of which are incorporated here by reference.

At least a portion of the work described was carried out under Cooperative Agreement NBCH2030002 from the United States Department of Interior with funding provided by the Department of Defense, and the United States has certain rights in this work.

BACKGROUND

This description relates to managing and using shared digital information on a network.

Most systems that share digital information on a network use properties of specific operating systems to control the access of users to information stored as files. Normally, certain parameters can be set to allow access, at one extreme, only to the user who is setting the parameters or, at the other extreme, to anyone without restriction.

Some operating systems allow definition of a group of members to whom access is allowed according to group permissions set by the owner of the stored information. Typically, a member of the group would log into the network-based information system by using a group name and a single private password assigned to the group. The group members are expected to keep the group name and password private in order not to compromise the information. In such as system, a user of a Machine A that is remote from a Machine B on which the information is stored cannot access the information without first logging into Machine B.

Another technique would be for all such potential users and uses to be made known to a third machine, Machine C, which could then be queried by Machine B to confirm that users on Machine A are permitted to make use of the information stored on Machine B.

Because not all users, for example a group of Internet users, of the shared information may be logged onto a single machine that controls access and use, an approach is needed for machine independent validation of users and groups of users

when, for example, a user on Machine A wants to access, modify, append, delete or perform other actions on information on Machine B.

SUMMARY

In general, in an aspect, use of and access to shared information on a network is managed by validating that a putative persistent identifier asserted to be associated with a user who is requesting to use information on a network is actually associated with the user, resolving the persistent identifier to state information about the user and a group to which the user belongs and that is associated with the identifier, determining whether the state information for the group contains the identifier associated with the user, and allowing access and/or use of the shared information based on the permissions allowed for the group.

Implementations may include one or more of the following features. There are two or more groups of users for whom use of the information is authorized. The user may be a person, a computer program, or a computerized process. The state information about the group indicates permissions or other uses authorized for members of that group. The group comprises only one authorized user and the identifier of the single user for that group is contained in the state information for the object. An identifier of a second group is contained in the state information of the first group and is used to allow use by a member of the second group of information that the first group is authorized to access. Aggregation in a digital object, of multiple inputs submitted by unrelated users to make assertions about the shared information, is enabled by associating an identifier with a metadata record that contains an identifier of each of the multiple inputs, and identifying the aggregate of the inputs using an identifier, in the state information, that is computed from an identifier of the shared information by addition of information that allows direct resolution and mapping from the original shared information's identifier to the identifier for the aggregate collection. The additional information is a set of characters added to the original object's identifier. The multiple sets correspond to a given digital object. The sets comprise metadata records. The sets provide commentary on information submitted by another party. The commentary includes processes that can be used to process the digital object. The state information is protected by a security system. The security system includes a public key infrastructure. The security system includes password protection. Metadata is searched across multiple metadata

registries; each of the set of registries makes available metadata to at least one registry; the one registry, using algorithms and/or computational procedures to map search requests into queries across different metadata schemas and/or controlled vocabularies used by the other registries in the set of multiple registries, the queries being used to progress the search and discovery process. In some examples, the one registry invokes a distributed search algorithm across the multiple registries. Elements of the one registry are distributed among the set of multiple metadata registries.

In general, in an aspect, a unique persistent identifier of a person or group is received in connection with a request by the person or group to perform an action with respect to a stored record, the unique persistent identifier resolved to state information, and, the state information is used to determine whether the person or group has authority to perform the action with respect to the stored record.

Other aspects include other combinations of the features and aspects described above and other features and aspects, described as methods, apparatus, compositions, processes, systems, program products, means for performing functions, and in other ways.

Other features and advantages will become apparent from the description and the claims.

DESCRIPTION

Figures 1 through 7 are schematic level block diagrams of aspects of registry systems.

As shown in figure 1, we describe here a way to ensure that the rights of authorized parties 10 (individuals or other persons or groups of such persons, as well as computer processes or running programs acting on behalf of such persons or groups of persons, sometimes called users) to modify, access, append to, delete, or perform other actions on (we sometimes refer to these actions and others by the general term "use") information 12 (e.g., digital objects 18) that is shared on a network 14, can be controlled independently of the locations 22 of the stored information or of the choice of platform (e.g., operating system) used to make it available to users. The authority of a user or even a group of users to work with the shared information is regulated based on the use of persistent machine-independent identifiers 16 that can be validated

quickly and automatically on the network without a need for human involvement. In addition, the techniques described here enable aggregating and appending information supplied by individuals who are members of at least one authorized group. The technique described here need not rely on a single machine to check authorizations.

The persistent identifiers that regulate the use of the shared information map 17 into state information 18, controlled by, for example, the owner 20 of the information. We sometimes refer to state information as metadata, metadata records, or metadata instances. Typically, shared information in the form of digital objects 18 will be held in repositories 22 and metadata registries 24 will assist users in discovering the digital information and helping owners 20 and administrators 30 of such information, including groups 40, 42, to manage permissions 41, 43 using a network based authentication and authorization mechanism.

The metadata registries 24 (which we sometimes call simply registries) hold information (which we sometimes call structured metadata records 26) about collections of information 12 (e.g., digital objects), such as collections contained in a set of repositories maintained by a user or a group. Users of a metadata registry may choose to or be required to register certain digital objects that belong to the collections with the registry.

The designation of digital objects and the repositories holding them is accomplished through a set of procedures that yield a set of identified repositories whose administrators 30 are authorized to deposit metadata records into a metadata registry and to maintain those records as an up-to-date reflection of the information 12 stored in their repositories. Each repository will, in general, have one or more managers whom we call repository managers or RepoMgrs31.

A RepoMgr will usually identify the individuals and groups (the administrators) that have rights to deposit and edit metadata records in the registry. Those individuals and groups will be given identifiers 32 and passwords 34 that may be used for identification and authentication. Authentication will ordinarily be transparent to an administrator, if the administrator has the necessary permissions; otherwise, the adminstrator will be denied the ability to carry out his requested task.

Many of the interactions with the registry will be carried out by highly automated administrator processes 36 not directly associated with any user or group. Those processes will usually have their own identities, for management purposes, but

they can equally well be assigned to identify the RepoMgr. In either case, the RepoMgr would bear responsibility for their actions. In general, a RepoMgr has responsibility for an entire collection of metadata records, while the term administrator is normally used to denote the individual, program or process that has responsibility for a particular metadata record in the collection. Depending on the role being played, a RepoMgr can assume the role of administrator in certain cases. Each identified administrator may also be a member of a group. All rights by an administrator to access and use a metadata record will be a combination of the rights associated with the administrator's group and, in the case of edits and deletions, the rights associated with the group that deposited the record

The right to insert can be authorized for specific administrative persons or groups independently of other rights. The RepoMgr can define multiple groups 40, 42, such that, for example, members of one group 40 can insert but not update or delete while members of a second group 42 can insert as well as update and delete. The update and delete actions will usually only be permitted to be done by members of the group that deposited the metadata record or of a related group, as described below. There can also be a special group of administrators (called the registry administration 45) who will have privileges over all records in the registry. This 'super-user' capability will allow the members of this group, for example, to untangle permission problems that may arise.

The RepoMgr for each repository may organize its users in groups having respective specified permissions as needed or into only a single group, if appropriate. This grouping and permission information is then communicated to the registry administration and enforced by authentication and authorization mechanisms 48 built into the registry. Tools can be used to simplify this task for the administrators.

As shown in figure 2, the metadata registries 50, 52 can be arranged to be used for generalized searching to find information based on attributes such as keywords, titles or author in addition to being used for rapid resolution of specific identifiers. The metadata records 64 in such registries can describe attributes 62 of the digital objects stored in the repositories, which are typically external to the registries.

Registry services 66, whether for searching or for rapid resolution, can aid in the discovery and location of remote digital objects using search tools. These tools allow users to query the collection of metadata records for identifiers or other

descriptions or attributes of objects that meet specified criteria. These search tools can be provided by the metadata registries or by external services that have access to the registries. The metadata registries may use standard metadata schemas that aid searching by providing consistent approaches to describing a set of digital objects, e.g., by having a consistently labeled portion of the metadata record describe the creator(s) of the relevant object. A metadata schema is a way to describe a class of information (e.g. books) and usually involves a description of a particular structure and taxonomy for the schema.

A single digital object can be the subject of multiple metadata records. This multiplicity of records may reflect different views on the underlying information, e.g., as described by different user communities, or reviews or annotations on the quality of the object. In other cases, they may detail the use of the object in other contexts, such as by supplying programs or identifiers for programs that can be used to transform or otherwise process the object in various ways.

In many cases, it is desirable to federate multiple metadata registries, such that a single query is executed against many registries. This federation 70 must take account of the fact that different registries will likely have different ways (metadata schemas) to describe their information to be useful to the community of users for which the registry was designed.

Building metadata registration systems that provide effective services across multiple registries, where each registry can contain multiple metadata records for the same object, more than one registry may contain metadata records for a given object, and the multiple registries use different metadata schemas, poses challenges.

Multiple metadata records registered in a single registry for the same digital object must be managed at one level as a single entity, e.g., in returning search results to users, and at other levels must be treated as separate entities, e.g., with respect to ownership of the metadata record and the ability to edit or delete incorrect records. Even if the records in a single registry use a single metadata schema, they still need to be aggregated. In practice, a single registry may support multiple metadata schemas.

The use of persistent identifiers and a rapid resolution system that maps identifiers into state information about the objects simplifies the coordination and use of multiple registries with multiple metadata schemas.

As shown in figure 3, each piece of information (e.g., a digital object 70) that is registered in the information system has a globally unique persistent identifier 72 that is registered within an identifier system 74, for example, the Global Handle Registry that is maintained as part of the Handle System managed by the Corporation for National Research Initiatives. This identifier 72 can later be resolved through a resolution service 76 to obtain indirection information 78, such as the location of the repository 80 or other service where the object may be found. Other current information about the object, such as which individuals or groups may have access to it 82, may also be kept in the identifier system.

The information kept in the identifier system for a given object tends to be the type of metadata used to access the object and which requires rapid resolution. The information kept in more general metadata registries tends to be of the type used to discover the object in the first place. For example, specifications for the Handle System and its component parts (see, for example, RFCs 3650, 3651, 3652) are incorporated here by reference. Also incorporated by reference is United States Patent 6,135,646 that was issued on October 24, 2000, to Robert E. Kahn and David K. Ely, which contains information about globally unique identifiers for digital objects, identifier systems, and identifier resolution systems, among other things.

The owner (or authorized holder) of an object will normally arrange to have an identifier assigned to it, but in some cases a metadata registry or a repository, by prior arrangement, may serve as the object owner's agent and manage the assignment of the identifier to the object and its registration in the identifier system or other rapid resolution system.

By a digital object, we mean, for example, the information to be made available from a given repository; which is the subject of the registration in the registry. Each digital object has a unique identifier. By a metadata record, we mean, for example, the metadata about a given digital object that is submitted to and indexed and stored by the registry.

Referring to figure 4, in a metadata registry 52, each metadata record 64 is itself assigned an identifier 84. Assignment and management of these identifiers are the responsibility of the owners 20 of the digital objects 18 and the administrators 98 of the metadata registry. Each set of metadata records referencing a given digital object is aggregated into a super metadata record 100 containing the identifiers 102 of all its

component records, and that super metadata record is itself assigned an identifier 104. The relationship among the related metadata records is maintained either directly in the identifier system 74 or in a data structure (the super metadata record) held within the metadata registry. In either case, having the identifier for any of the related metadata records allows a user or a process to find all of the related records. If held in the identifier system, the first level metadata records can lead directly to the super metadata record, which contains references to the identifiers for all of the related metadata records. If the relationship is maintained in a data structure in the metadata registry, then the identifier for each of the first level metadata records will lead to that structure or service in that metadata registry.

The authority to administer 108 each metadata record is maintained by the registry and the registry allowsr organizations 110 and other users, some of which will be structured hierarchically, to administer the records and manage group administration 112. Each metadata record is also maintained by the registry together with information about who owns, controls or administers the record 114 and, thus, who can create and register new metadata records, can edit metadata records, and can delete metadata records. The information can also designate which other users can aggregate related information and make it available in a given object's metadata record. Management of the use of the metadata record consistently with these designations can be achieved through the use of persistent identifiers and a rapid resolution mechanism.

As shown in figure 5, in some implementations of a basic control mechanism for submitting and managing metadata records, each submitter (e.g., a person or a process) 118 is assigned an individual identifier 120. We sometimes use the word submitter to refer to a user whose activities relate to the creation, submission, and management of metadata records 122 in the metadata registry 52, to distinguish such a person from a user who is making use of the registry to find and use the digital objects held in the repositories. The term "user" is a general term in this context, and a "submitter" is a type of user.

Each submission 124 of one or more metadata records 122 to the metadata registry 52 contains, among other data items, the identifier 120 for the submitter 118 making the submission and an identifier 126 for a defined group of which the submitter is a member. In some cases, an additional service, such as an existing authentication service128, holds the submitters' identifiers, and the submission 124 containing the

metadata records also contains authentication data 130, e.g., a password, which the registry uses to obtain the submitter's actual identifier, e.g., a password 132.

If the metadata submission contains the submitter's identifier 120, and the registry wishes to authenticate that the submitter is who he claims to be, the registry can authenticate the submitter by the commonly used challenge/response technique using public key technology. The identifier system can be used to provide, in a trusted fashion, the Public Key Infrastructure (PKI) used to store and make available the public key for any submitter designated with an identifier. Other PKI approaches may be used. The submission may contain other authentication information using alternate authentication approaches; in this latter case, the registry uses those alternate approaches to obtain the identifier of the submitter with the requisite assurances of that approach to assure that the submitter is who he claims to be to whatever level of assurance is set by the policies for a given metadata registry.

For group permissions, each submitter is assigned to be part of one or more groups 140 of submitters. Each group is identified by an identifier 126. Each submission 124 also contains the identifier 126 for the group under the aegis of which the submitter is making the submission. The identifier record 129 for each of these groups holds the complete list 142 of the identifiers for submitters who are part of that group. After discovering or otherwise authenticating the identifier for each submitter, the metadata registry will resolve 144 the group identifier to verify that the user is a member of the group to which the submitter claims to belong.

At this point in processing the metadata record submission, the registry has identified the submitter, authenticated (or not) the submitter based its identifier (according to its policies), and, if required, validated that the submitter is or is not a member of the group under whose aegis the submission is being made.

The registry now determines if the submitter, under the aegis of the given group, has the rights to perform the requested actions, as identified 146 in the submission. Each submission to a metadata registry is a request for an action. The most common action is the submission of a new metadata record, but other actions are possible. A given metadata record may have to be edited, for example, because the initial submission was incorrect or the current state of the described digital object has changed. In some cases, if an object is no longer available or if the owner or authorized holder of the object no longer wishes it to be discovered, it may be necessary to delete

(or otherwise make inaccessible) the metadata record. In general, the permissions to perform these different actions will vary according to individual and group.

The registry performs the validation by resolving 152 the identifier 150 for the registry itself, which produces the set of permissions 156 each group holds relative to a set of digital objects. The set of objects 154 are referenced by the identifier prefix. The permissions 156 are shown in Figure 5 as single characters, e.g. "i" for insert and "u" for update, etc.

When an object that is referenced in the registry is deleted, the registry is responsible for ensuring that its references continue to make sense. The repository administrator notifies the registry of such deletions.

All submitter rights and related permissions can be held in the identifier system 74 and thus all rights calculations can be performed using data obtained through resolution of identifiers by the identifier system. The operational requests of authenticated submitters, e.g., to create a new metadata instance (MI), are granted or denied after matching the submitter authorization against the requested operation. In some efficient implementations of this system, all authorization is done at the group level and not at the level of an individual submitter, where the submitter is part of a group and is so authorized by the group. For example, an individual submitter would be part of a group consisting of one member. However, other implementations are also possible, such as one in which an individual submitter can be handled directly if the system tests for both individual submitters and for group submitters.

There are two basic sets of permissions, one relatively complex set associated with the registry as a whole and one simple type associated with individual MIs. In some cases, each MI is associated with the group that created it. For example, the registry as a whole could have the following eight separate permissions, each turned on or off at the level of the group:

Insert a new MI into the registry.

Delete an MI created by the given group.

Update an MI created by the given group.

De-activate an MI created by the given group (i.e. remove it from search indices, but do not actually delete it).

Activate an MI previously de-activated (i.e., put it back in the search indices).

Read Metadata: reserved, but intended to apply to secure records.

Read Index: reserved, as above.

Identifier operations: records the fact that the given group was allocated a given set of identifier attributes (e.g., Handle System prefixes), and so the group can request the registry, which is serving as administrator for those prefixes, to insert or edit identifiers under those prefixes).

The following example illustrates this permissions mechanism.

Consider a hypothetical group of submitters associated with an organization A.B.C. Assume their identifiers are denoted as A.B.C/groucho, A.B.C/harpo, and A.B.C/chico. They are grouped into A.B.C/marxbros, which is their group identifier. That group identifier may be used to obtain all of the individual identifiers for members of the group. Permissions, such as the ability to insert a new MI into a given metadata registry are accessed using the group identifier. In the case of an insertion, this involves associating each relevant group with the identifier record 150 for the registry and recording which permissions apply to each group with respect to the registry.

Because listing every group directly under the single identifier for a registry will not scale for very large registries, a further level of indirection may be used. In that case, each registry maps to multiple identifiers, each of which is used to reference a set of group permission records. The relevant metadata record for a given group identifier is discovered by resolving the group identifier. So if A.B.C/marxbros has permission to insert information, but not permission to delete information, then each of Groucho, Harpo, and Chico can insert a new metadata record into the given registry but could never delete it. If a new individual, e.g., zeppo, was added to the group then A.B.C/zeppo would be added to A.B.C/marxbros and zeppo would also be able to insert but would not be able to delete, and so on.

Extending our example, groucho believes a given insertion, (e.g. the metadata record identified as 100.1000.1/asdfg) to be a hoax and attempts to delete the record. This returns an error message, because even though groucho is validated as a member of the group that created the record, that group itself does not have delete permissions on the registry.

If the record was created by A.B.C/marxbros and A.B.C/marxbros does not have delete permission, how can the metadata record corresponding to 100.1000.1/asdfg ever be deleted? As shown in figure 4, the answer is that groups can be nested, so that a group 172 can be a member of another group 170. This can be done at the individual organization level 110, and it can also be done across all organizations

to create a super group 176 for a given registry. Thus, one can enforce a rule stating that only members of the group that inserted a given metadata record can remove or edit that record, while still grouping submitters into different permission categories. In our example, suppose a second group, identified as A.B.C/librarians,initially consists of a single member, A.B.C/Marian, will have delete permission, and is added as a member of A.B.C/marxbros. Marian can now successfully submit a delete request for 100.1000.1/asdfg as a member of A.B.C/librarians, because the system would Authenticate Marian as A.B.C/Marian; determine that Marian is part of A.B.C/librarians, as she claims; determine that the group that created 100.1000.1/asdfg is A.B.C/marxbros; determine that A.B.C/librarians is part of A.B.C/marxbros (in some cases this could be nested down several layers), and determine that A.B.C/librarians has delete rights. While this mechanism could be used to create an unlimited hierarchical set of permissions, it will generally be restricted to two levels.

A given organization can thus divide its administrators into two or more groups, with some members having oversight privileges over other members' submitting and editing data.

The same mechanism is used to determine who can append information to a digital object. An individual may be designated as having permission to append information, even if he cannot delete or modify the object or other parts of the metadata record. Normally, such an individual would have the ability to access the object, and it is possible to allow such individual to append information even though he cannot access the object. This might be useful for administrators who maintain the permissions system, but who are themselves not permitted to access the information.

Similarly, the same mechanism is used to determine who can aggregate information for inclusion in a metadata record. If, for example, several individuals append information that is deemed to have common attributes by a party with permission to aggregate, that party may modify the metadata record to remove individual appendages and reappend them as a collection of appendages. This may be useful, for example, in categorizing types of appendages, or for consolidating proposed edits of a given document by a group of collaborators or reviewers.

As shown in figures 6 and 7, multiple metadata registries 200, 202 can be federated by permitting creation of federation level metadata 204, 206 specifically for formation of registries of registries 208 (as in a hierarchy of registries) or in the

construction and execution of a class of distributed query mechanisms that can be used efficiently with multiple registries without the need to form registries of registries. We do not here describe how to enable cross-schema indexing or cross-domain searching. Instead, an infrastructure approach is described that can accommodate multiple solutions including those mentioned above.

Figure 6 illustrates the approach to aggregating data from the multiple metadata registries 200, 202 into a single registry of registries 208. A single federation of registries is illustrated, but the same approach can be applied to federating multiple registries of registries such that each registry of registries could itself be considered a single registry; and a new registry of such registries could further federate them, resulting in an even larger federation. Each registry (as mentioned earlier) and each registry of registries is identified by a persistent identifier 210. The registry of registries enables queries or provides other services 212 across multiple registries as though they constituted a single registry. Each registry of registries holds the relevant data 220 from the individual registries that enable these federated services.

The specifics of that federation data depends on the domain in which the registries operate, the services being offered, and the specific metadata schemas being used for describing the registered content. In some examples, the registry of registries could contain, the information about which of its constituent metadata registries use which metadata schemas 222 and provides a set of translation facilities 224 between schemas, allowing it to take a single query and map it across the multiple registries. This translation facility could use keyword substitution, ontologies, or other semantic mapping strategies. Other federation data aggregated at a registry of registries could include access control policies 226, subject level coverage, and perhaps even technical details of how to prepare efficient queries.

Some federation data collected at a given registry of registries will apply to the constituent metadata registries individually, e.g., registry X has one set of access policies and registry Y has another. That is, the data applies to the metadata registry as a whole and does not vary across the individual metadata records contained in the registry. Other federation data, collected by a registry of registries, however, will apply at the individual repository or metadata record level. For example, a community of registries could agree to allow its constituent metadata registries to use multiple metadata schemas, where each registry uses the schema specific to its users, but also

agrees upon a schema that applies across the entire set of metadata registries. The schema used across all registrations in all of the registries in a given collection of registries would be more general, for example, in subject description terms, than the individual registry level schema, and could be used to determine which individual registries were worth querying at a more detailed level. This would require that each metadata record submitted to each of the registries within a given registry of registries community would be submitted with metadata created according to two different schemas. The first detailed set of metadata formed according to the specific schema would be kept at each individual registry for use in queries directed at that registry while the second set of metadata formed according to the more general schema would be forwarded on to the registry of registries. A query executed against that second and more general set of data would show which registries contained records worth searching in a more detailed fashion.

Figure 7 illustrates a submission 230 of metadata records 232 containing multiple levels, including record level metadata 234, registry level metadata 236, and registry of registries level metadata 238.

Registries can be arbitrarily grouped across common sets of characteristics, such as their service interfaces, taxonomies, metadata schemas, and data collections. Each such group may define its own criteria and characteristics and interfaces. Each individual group is queried through a higher level registry, known as an Aggregator, which interacts with one or more of the registries in each group and acquires the characteristics of that group and generates a response that is consistent with the interface characteristics of that group.

The identifier maps to values that identify the characteristics of the groups and the services of the individual registries. From these values, a deterministic map may be created between the groups and the characteristics of the groups. For example, an Aggregator talking to two registries in one group, but each with different schemas and taxonomies (e.g. one schema is Dublin Core and the other is based on MARC records) would extract a predefined path through the use of the identifiers to actually do the conversions. The method is completely general and the result is thus deterministically derived from the predefined information.

The Aggregator service is able to bridge the heterogeneity of the various registries in the different groups by resolving each of the group identifiers into their

respective characteristics. From these identified characteristics, the Aggregator can deterministically compute effective queries, and related data conversions by mapping the identifiers of the group characteristics to that of the individual registries in order to generate a consistent compound result.

Objects and the metadata associated with them can be represented in a way that enables unrelated parties to make assertions about a particular object previously registered by its owner or responsible party. The registry provides an abstract digital object representation mechanism. This abstract representation is also expressed as a digital object. The representation groups a set of metadata instance assertions about the digital object. Therefore, each set of metadata records referencing the same digital object is aggregated into a super metadata record consisting of the identifiers of all its component records, and that super metadata record is itself assigned an identifier. The relationship among the related metadata records is maintained either directly in the identifier system or in a data structure held within the metadata registry. In either case, having the identifier for any of the related metadata records allows a process or other user to find all of the related records. If held in the identifier system, the first level metadata records can lead directly to the super metadata record, which contains references to the identifiers for all of the related metadata records. If the relationship is maintained in a data structure in the metadata registry, then the identifier for each of the first level metadata records will lead to that structure or service in that metadata registry.

The registry is able to perform authenticated actions over the object representation and the metadata instances in a distributed independent fashion. This method enables identifiers for metadata instances to be delegated to and maintained by the registry which is responsible for user authentication. The method allows users to operate over their particular metadata instances, and the future aggregation of these instances across multiple systems by means of the expressed object metadata instance relationship.

Other embodiments are within the scope of the following claims.

WHAT IS CLAIMED IS:

1. A method to manage access to and use of shared information on a network comprising

validating that a putative persistent identifier asserted to be associated with a user who is requesting to use information on a network is actually associated with the user,

resolving the persistent identifier to state information about the user and a group to which the user belongs and that is associated with the identifier,

determining whether the state information for the group contains the identifier associated with the user and

allowing access and/or use of the shared information based on the permissions allowed for the group.

2. The method of claim 1 in which there are two or more groups of users for whom use of the information is authorized.

3. The method of claim 1 in which the user may be comprise a person, computer program or computerized process.

4. The method of claim 1 in which the state information about the group indicates permissions and other uses authorized for members of that group.

5. The method of claim 1, 2, 3, or 4 in which the group comprises only one authorized user and the identifier of the single user for that group is contained in the state information for the object.

6. The method of claim 1 in which an identifier of a second group is contained in the state information of the first group and is used to allow use by a member of the second group to information that the first group is authorized to access.

7. The method of claim 1 also comprising

enabling aggregation in a digital object, of multiple inputs submitted by unrelated users to make assertions about the shared information, by

associating an identifier with a metadata record that contains an identifier of each of the multiple inputs, and

identifying the aggregate of the inputs using an identifier, in the state information, that is computed from an identifier of the shared information by addition

of information that allows direct resolution and mapping from the original shared information's identifier to the identifier for the aggregate collection.

8. The method of claim 7 in which the additional information is a set of characters added to the original object's identifier.

9. The method of claim 7 in which the multiple sets all correspond to a given digital object.

10. The method of claim 7 in which the sets comprise metadata records.

11. The method of claim 7 in which the sets provide commentary on information submitted by another party.

12. The method of claim 11 in which the commentary includes processes that can be used to process the digital object.

13. The method of claim 1 or 7 in which the state information is protected by a security system.

14. The method of claim 13 in which the security system comprises a public key infrastructure.

15. The method of claim 13 in which the security system comprises password protection.

16. The method of claim 1 or 7 comprising

searching metadata across multiple metadata registries, each of the set of registries making available metadata to at least one aggregator,

the aggregator using algorithms and/or computational procedures to map search requests into queries across different metadata schemas and/or controlled vocabularies used by the other registries in the set of multiple registries, and

the queries being used to progress the search and discovery.

17. The method of claim 16 in which the aggregator invokes a distributed search algorithm across the multiple registries,

18. The method of claim 16 or 17 in which elements of the aggregator are distributed among the set of multiple metadata registries.

19. The method of Claim 1 comprising the identifying of registries and groups of registries so as to enable access to their defined respective service interfaces, taxonomies, metadata schemas, and data collections.

20. The method of Claim 19 in which the identifiers are resolved to obtain their respective characteristics and attributes and the attributes are then combined

deterministically into a unique path that can be uniquely resolved into the appropriate conversion utility.

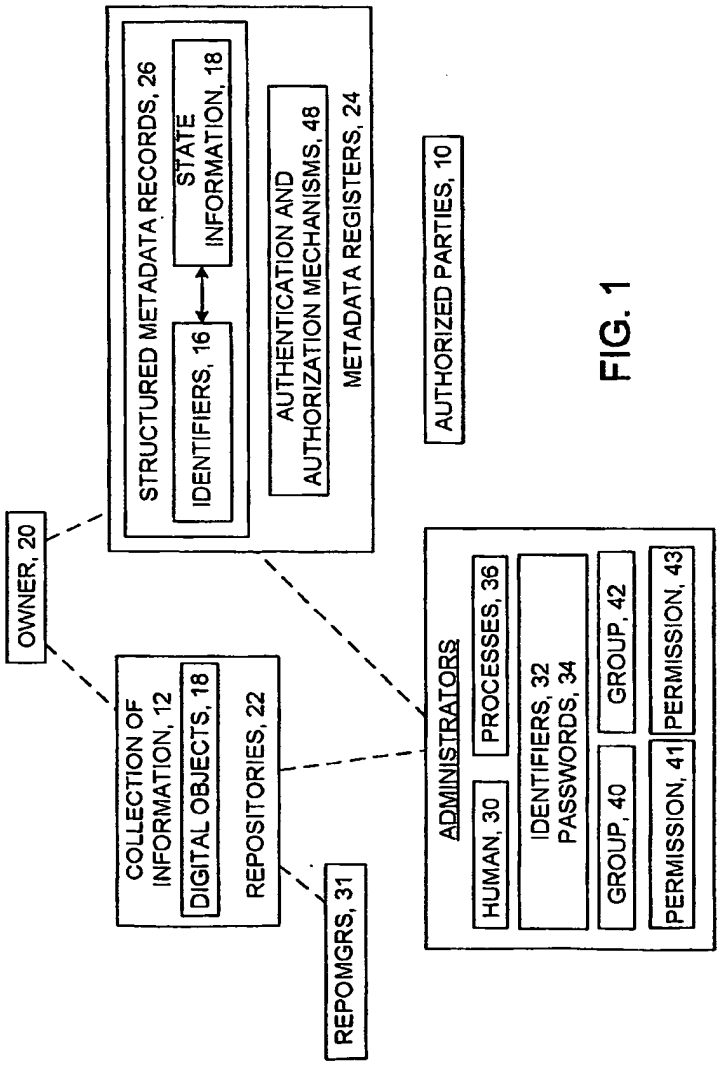
STRUCTURED METADATA RECORDS, 26
IDENTIFIERS, 16
STATE INFORMATION, 18
AUTHENTICATION AND AUTHORIZATION MECHANISMS, 48
METADATA REGISTERS, 24
AUTHORIZED PARTIES, 10
OWNER, 20
COLLECTION OF INFORMATION, 12
DIGITAL OBJECTS, 18
REPOSITORIES, 22
REPOMGRS, 31
ADMINISTRATORS
HUMAN, 30
PROCESSES, 36
IDENTIFIERS, 32
PASSWORDS, 34
GROUP, 40
GROUP, 42
PERMISSION, 41
PERMISSION, 43

FIG. 1

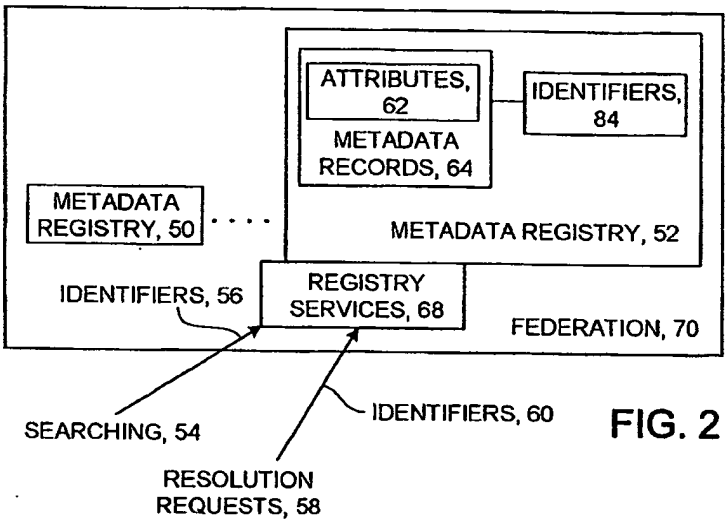
ATTRIBUTES, 62
IDENTIFIERS, 84
METADATA RECORDS, 64
METADATA REGISTRY, 50
METADATA REGISTRY, 52
IDENTIFIERS, 56
REGISTRY SERVICES, 68
FEDERATION, 70
SEARCHING, 54
IDENTIFIERS, 60
RESOLUTION REQUESTS, 58

FIG. 2

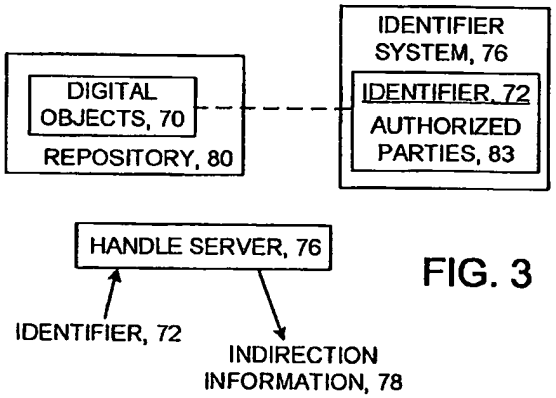
IDENTIFIER SYSTEM, 76
DIGITAL OBJECTS, 70
IDENTIFIER, 72
AUTHORIZED PARTIES, 83
REPOSITORY, 80
HANDLE SERVER, 76
IDENTIFIER, 72
INDIRECTION INFORMATION, 78

FIG. 3

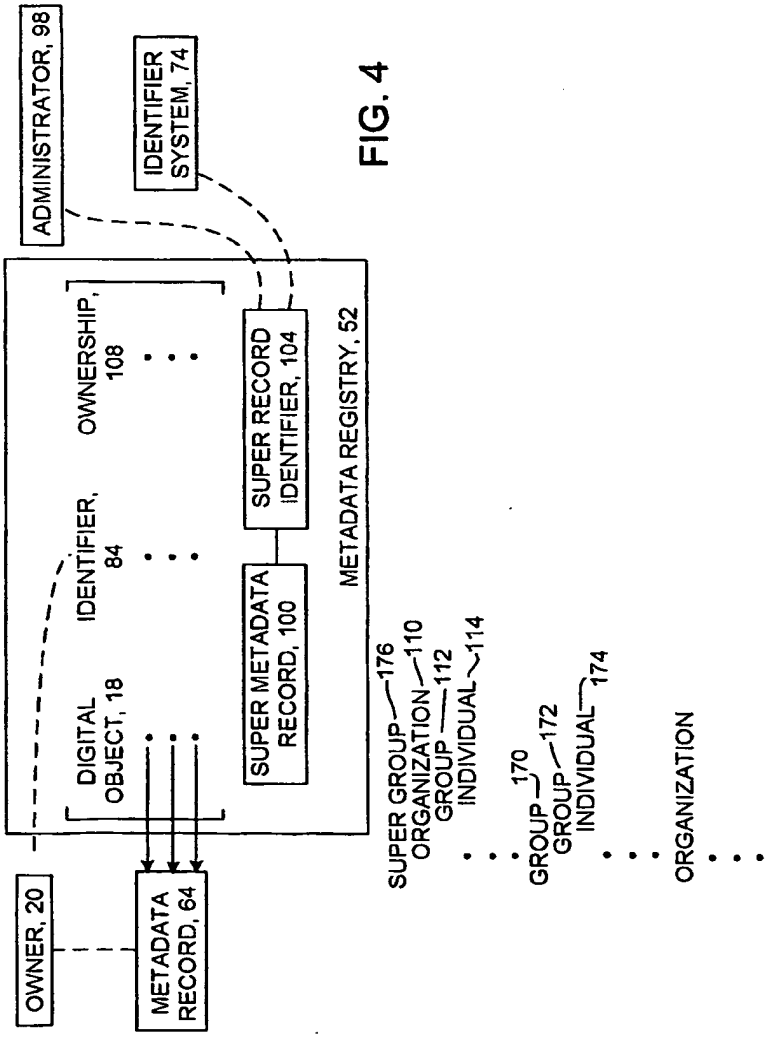
ADMINISTRATOR, 98
IDENTIFIER SYSTEM, 74
METADATA REGISTRY, 52
DIGITAL OBJECT, 18
IDENTIFIER, 84
OWNERSHIP, 108
SUPER METADATA RECORD, 100
SUPER RECORD IDENTIFIER, 104
OWNER, 20
METADATA RECORD, 64
SUPER GROUP 176
ORGANIZATION 110
GROUP 112
INDIVIDUAL 114
GROUP 170
GROUP 172
INDIVIDUAL 174
ORGANIZATION

FIG. 4

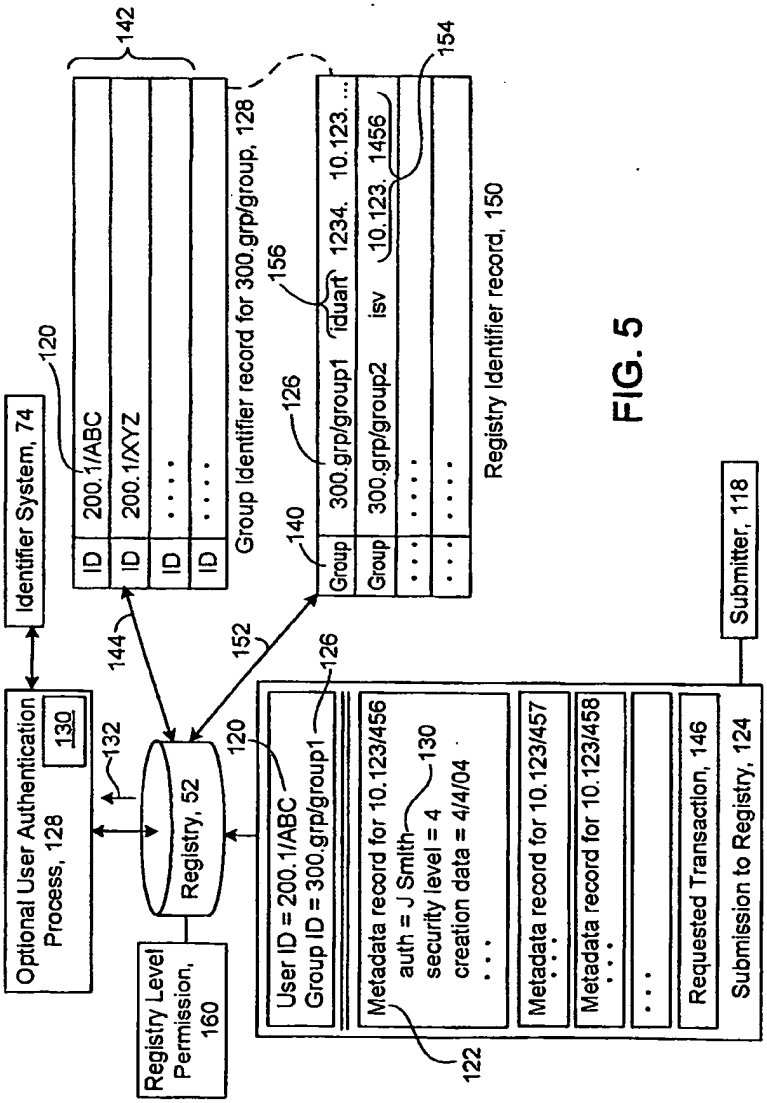
Identifier System, 74
120
Optional User Authentication Process, 128
130
132
Registry, 52
Registry Level Permission, 160
144
152
142
ID 200.1/ABC
ID 200.1/XYZ
ID . . . .
ID . . . .
Group Identifier record for 300.grp/group, 128
140
Group 300.grp/group1 iduart 1234. 10.123. ...
Group 300.grp/group2 isv 10.123. 1456
. . . . . .
. . . . . .
Registry Identifier record, 150
154
156
126
User ID = 200.1/ABC
Group ID = 300.grp/group1
Metadata record for 10.123/456
auth = J Smith
security level = 4
creation data = 4/4/04
. . .
Metadata record for 10.123/457
Metadata record for 10.123/458
. . .
Requested Transaction, 146
Submission to Registry, 124
Submitter, 118
122

FIG. 5

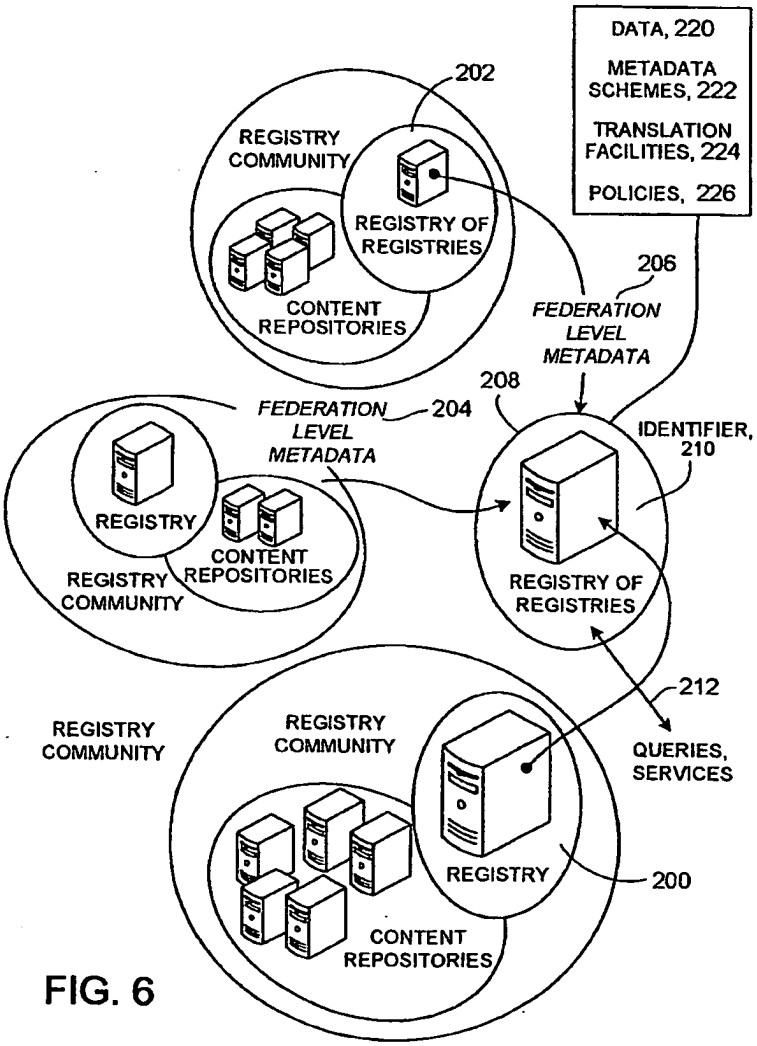
DATA, 220
METADATA SCHEMES, 222
TRANSLATION FACILITIES, 224
POLICIES, 226
202
REGISTRY COMMUNITY
REGISTRY OF REGISTRIES
CONTENT REPOSITORIES
206
FEDERATION LEVEL METADATA
208
FEDERATION LEVEL METADATA
204
IDENTIFIER, 210
REGISTRY
CONTENT REPOSITORIES
REGISTRY COMMUNITY
REGISTRY OF REGISTRIES
212
QUERIES, SERVICES
REGISTRY COMMUNITY
REGISTRY COMMUNITY
REGISTRY
200
CONTENT REPOSITORIES

FIG. 6

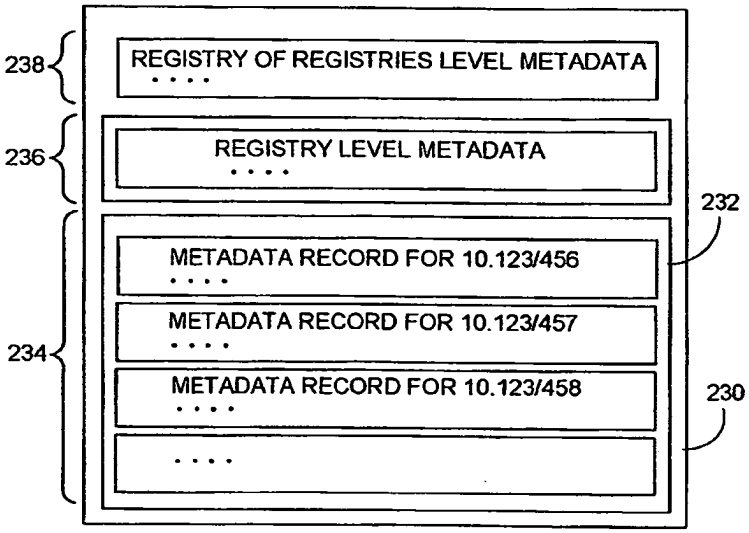
238
REGISTRY OF REGISTRIES LEVEL METADATA
236
REGISTRY LEVEL METADATA
232
234
METADATA RECORD FOR 10.123/456
METADATA RECORD FOR 10.123/457
METADATA RECORD FOR 10.123/458
230

FIG. 7